

Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: June 7, 2015

IJ. Wijnands, Ed.
Cisco Systems, Inc.
E. Rosen, Ed.
Juniper Networks, Inc.
A. Dolganow
Alcatel-Lucent
J. Tantsura
Ericsson
S. Aldrin
Huawei Technologies
December 4, 2014

Encapsulation for Bit Index Explicit Replication in MPLS Networks
draft-wijnands-mpls-bier-encapsulation-02

Abstract

Bit Index Explicit Replication (BIER) is an architecture that provides optimal multicast forwarding through a "multicast domain", without requiring intermediate routers to maintain any per-flow state or to engage in an explicit tree-building protocol. When a multicast data packet enters the domain, the ingress router determines the set of egress routers to which the packet needs to be sent. The ingress router then encapsulates the packet in a BIER header. The BIER header contains a bitstring in which each bit represents exactly one egress router in the domain; to forward the packet to a given set of egress routers, the bits corresponding to those routers are set in the BIER header. The details of the encapsulation depend on the type of network used to realize the multicast domain. This document specifies the BIER encapsulation to be used in an MPLS network.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on June 7, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
2. The BIER-MPLS Label	3
3. BIER Header	5
4. Imposing and Processing the BIER Encapsulation	8
5. IANA Considerations	10
6. Security Considerations	10
7. Acknowledgements	10
8. Contributor Addresses	10
9. References	12
9.1. Normative References	12
9.2. Informative References	12
Authors' Addresses	12

1. Introduction

[BIER_ARCH] describes a new architecture for the forwarding of multicast data packets. That architecture provides optimal forwarding of multicast data packets through a "multicast domain". However, it does not require any explicit tree-building protocol, and does not require intermediate nodes to maintain any per-flow state. That architecture is known as "Bit Index Explicit Replication" (BIER).

This document will use terminology defined in [BIER_ARCH].

A router that supports BIER is known as a "Bit-Forwarding Router" (BFR). A "BIER domain" is a connected set of Bit-Forwarding Routers (BFRs), each of which has been assigned a BFR-prefix. A BFR-prefix is a routable IP address of a BFR, and is used by BIER to identify a BFR. A packet enters a BIER domain at an ingress BFR (BFIR), and leaves the BIER domain at one or more egress BFRs (BFERs). As

specified in [BIER_ARCH], each BFR of a given BIER domain is provisioned to be in one or more "sub-domains". In the context of a given sub-domain, each BFIR and BFER must have a BFR-id that is unique within that sub-domain. A BFR-id is just a number in the range [1,65535] that, relative to a BIER sub-domain, identifies a BFR uniquely.

As described in [BIER_ARCH], BIER requires that multicast data packets be encapsulated with a header that provides the information needed to support the BIER forwarding procedures. This information includes the sub-domain to which the packet has been assigned, a Set-Id (SI), a BitString, and a BitStringLength. Together these values identify the set of BFERs to which the packet must be delivered.

This document is applicable when a given BIER domain is both an IGP domain and an MPLS network. In this environment, the BIER encapsulation consists of two components:

- o an MPLS label (which we will call the "BIER-MPLS label"); this label appears at the bottom of a packet's MPLS label stack.
- o a BIER header, as specified in Section 3.

Following the BIER header is the "payload". The payload may be an IPv4 packet, an IPv6 packet, an ethernet frame, or an MPLS packet. If it is an MPLS packet, then an MPLS label stack immediately follows the BIER header. The top label of this MPLS label stack may be either a downstream-assigned label ([RFC3032]) or an upstream-assigned label ([RFC5331]). The BIER header contains information identifying the type of the payload.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. The BIER-MPLS Label

As stated in [BIER_ARCH], when a BIER domain is also an IGP domain, IGP extensions can be used by each BFR to advertise the BFR-id and BFR-prefix. The extensions for OSPF are given in [OSPF_BIER_EXTENSIONS]. The extensions for ISIS are given in [ISIS_BIER_EXTENSIONS].

When a particular BIER domain is both an IGP domain and an MPLS network, we assume that each BFR will also use IGP extensions to advertise a set of one or more "BIER-MPLS" labels. When the domain contains a single sub-domain, a given BFR needs to advertise one such label for each combination of SI and BitStringLength. If the domain

contains multiple sub-domains, a BFR needs to advertise one such label per SI per BitStringLength for each sub-domain.

The BIER-MPLS labels are locally significant (i.e., unique only to the BFR that advertises them) downstream-assigned MPLS labels. For example, suppose that there is a single sub-domain (the default sub-domain), that the network is using a BitStringLength of 256, and that all BFRs in the sub-domain have BFR-ids in the range [1,512]. Since each BIER BitString is 256 bits long, this requires the use of two SIs: SI=0 and SI=1. So each BFR will advertise, via IGP extensions, two MPLS labels for BIER: one corresponding to SI=0 and one corresponding to SI=1. The advertisements of these labels will also bind each label to the default sub-domain and to the BitStringLength 256.

As another example, suppose a particular BIER domain contains 2 sub-domains (sub-domain 0 and sub-domain 1), supports 2 BitStringLengths (256 and 512), and contains 1024 BFRs. A BFR that is provisioned for both sub-domains, and that supports both BitStringLengths, would have to advertise the following set of BIER-MPLS labels:

- L1: corresponding to sub-domain 0, BitStringLength 256, SI 0.
- L2: corresponding to sub-domain 0, BitStringLength 256, SI 1.
- L3: corresponding to sub-domain 0, BitStringLength 256, SI 2.
- L4: corresponding to sub-domain 0, BitStringLength 256, SI 3.
- L5: corresponding to sub-domain 0, BitStringLength 512, SI 0.
- L6: corresponding to sub-domain 0, BitStringLength 512, SI 1.
- L7: corresponding to sub-domain 1, BitStringLength 256, SI 0.
- L8: corresponding to sub-domain 1, BitStringLength 256, SI 1.
- L9: corresponding to sub-domain 1, BitStringLength 256, SI 2.
- L10: corresponding to sub-domain 1, BitStringLength 256, SI 3.
- L11: corresponding to sub-domain 1, BitStringLength 512, SI 0.
- L12: corresponding to sub-domain 1, BitStringLength 512, SI 1.

The above example should not be taken as implying that the BFRs need to advertise 12 individual labels. For instance, instead of advertising a label for <sub-domain 1, BitStringLength 512, SI 0> and

a label for <sub-domain 1, BitStringLength 512, SI 1>, a BFR could advertise a contiguous range of labels (in this case, a range containing exactly two labels) corresponding to <sub-domain 1, BitStringLength 512>. The first label in the range could correspond to SI 0, and the second to SI 1. The precise mechanism for generating and forming the advertisements is outside the scope of this document. See [OSPF_BIER_EXTENSIONS] and [ISIS_BIER_EXTENSIONS].

Note that, in practice, labels only have to be assigned if they are going to be used. If a particular BIER domain supports BitStringLengths 256 and 512, but some sub-domain, say sub-domain 1, only uses BitStringLength 256, then it is not necessary to assign labels that correspond to the combination of sub-domain 1 and BitStringLength 512.

When a BFR receives an MPLS packet, and the next label to be processed is one of its BIER-MPLS labels, it will assume that a BIER header (see Section 3) immediately follows the stack. It will also infer the packet's sub-domain, SI, and BitStringLength from the label. The packet's "incoming TTL" (see below) is taken from the TTL field of the label stack entry that contains the BIER-MPLS label.

The BFR MUST perform the MPLS TTL processing correctly. If the packet is forwarded to one or more BFR adjacencies, the BIER-MPLS label carried by the forwarded packet MUST have a TTL field whose value is one less than that of the incoming TTL. (Of course, if the incoming TTL is 1, the packet will not be forwarded at all, but will be discarded as an MPLS packet whose TTL has been exceeded.)

3. BIER Header

The BIER header is shown in Figure 1. This header appears after the end of the MPLS label stack, immediately after the MPLS-BIER label.

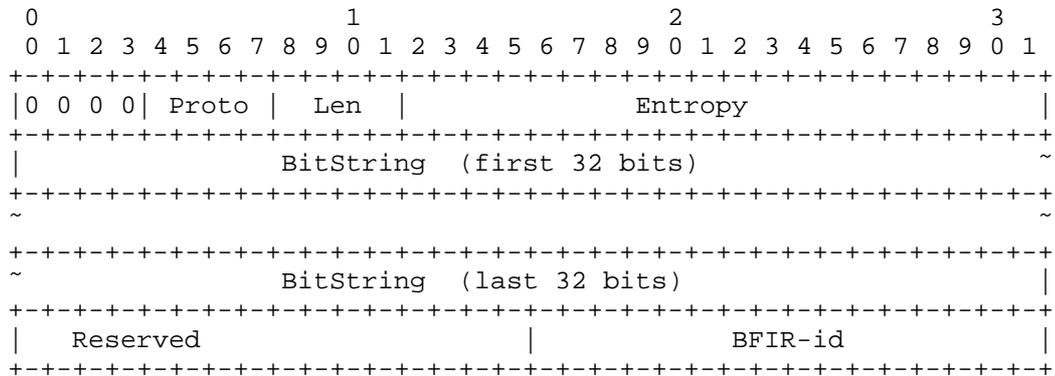


Figure 1: BIER Header

Ver:

The first 4 bits of the header are all set to zero; this ensures that the BIER header will not be confused with an IP header. This field can also be used as a version number if there are future revisions of the BIER header. However, the values 4 and 6 MUST NOT be used, as that may make the packets appear to some hardware forwarder to be IP packets.

Proto:

This 4-bit field identifies the type of the payload. (The "payload" is the packet or frame immediately following the BIER header.) The protocol field may take any of the following values:

- 1: MPLS packet with downstream-assigned label at top of stack.
- 2: MPLS packet with upstream-assigned label at top of stack (see [RFC5331]). If this value of the Proto field is used, the I bit MUST be set, and the BFR-id of the BFIR must be placed in the BFIR-id field. The BFIR-id provides the "context" in which the upstream-assigned label is interpreted.
- 3: Ethernet frame.
- 4: IPv4 packet.
- 6: IPv6 packet.

Len:

This 4-bit field encodes the length in bits of the BitString. If k is the length of the BitString, the value of this field is $\log_2(k)-5$. However, only certain values are supported:

- 1: 64 bits
- 2: 128 bits
- 3: 256 bits
- 4: 512 bits
- 5: 1024 bits
- 6: 2048 bits
- 7: 4096 bits

All other values of this field are illegal.

Entropy:

This 20-bit field specifies an "entropy" value that can be used for load balancing purposes. The BIER forwarding process may do equal cost load balancing, but the load balancing procedure MUST choose the same path for any two packets have the same entropy value.

If a BFIR is encapsulating (as the payload) MPLS packets that have entropy labels, the BFIR MUST ensure that if two such packets have the same MPLS entropy label, they also have the same value of the BIER entropy field.

BitString:

The BitString that, together with the packet's SI, identifies the destination BFERs for this packet. Note that the SI for the packet is inferred from the BIER-MPLS label that precedes the BIER header.

BFIR-id

By default, this is the BFR-id of the BFIR, in the sub-domain to which the packet has been assigned. The BFR-id is encoded in the 16-bit field as an unsigned integer in the range [1,65535].

Certain applications may require that the BFIR-id field contain the BFR-id of a BFR other than the BFIR. However, that usage of the BFIR-id field is outside the scope of the current document.

4. Imposing and Processing the BIER Encapsulation

When a BFIR receives a multicast packet from outside the BIER domain, the BFIR carries out the following procedure:

1. By consulting the "multicast flow layer" ([BIER_ARCH]), it determines the value of the "Proto" field.
2. By consulting the "multicast flow layer", it determines the set of BFERs that must receive the packet.
3. If more than one sub-domain is supported, the BFIR assigns the packet to a particular sub-domain. Procedures for determining the sub-domain to which a particular packet should be assigned are outside the scope of this document.
4. The BFIR looks up the BFR-id, in the given sub-domain, of each of those BFERs.
5. The BFIR converts each such BFR-id into (SI, BitString) format, as described in [BIER_ARCH].
6. All such BFR-ids that have the same SI can be encoded into the same BitString. Details of this encoding can be found in [BIER_ARCH]. For each distinct SI that occurs in the list of the packet's destination BFERs:
 - a. The BFIR makes a copy of the multicast data packet, and encapsulates the copy in a BIER header (see Section 3). The BIER header contains the BitString that represents all the destination BFERs whose BFR-ids (in the given sub-domain) correspond to the given SI. It also contains the BFIR's BFIR-id in the sub-domain to which the packet has been assigned.

N.B.: For certain applications, it may be necessary for the BFIR-id field to contain the BFR-id of a BFR other than the BFIR that is creating the header. Such uses are outside the scope of this document, but may be discussed in future revisions.

- b. The BFIR then applies to that copy the forwarding procedure of [BIER_ARCH]. This may result in one or more copies of

the packet (possibly with a modified BitString) being transmitted to a neighboring BFR.

- c. Before transmitting a copy of the packet to a neighboring BFR, the BFR finds the BIER-MPLS label that was advertised by the neighbor as corresponding to the given SI, sub-domain, and BitStringLength. An MPLS label stack is then prepended to the packet. This label stack [RFC3032] will contain one label, the aforementioned BIER-MPLS label. The "S" bit MUST be set, indicating the end of the MPLS label stack. The TTL field of this label stack entry is set according to policy. The packet may then be transmitted to the neighboring BFR. (This may result in additional MPLS labels being pushed on the stack. For example, if an RSVP-TE tunnel is used to transmit packets to the neighbor, a label representing that tunnel would be pushed onto the stack.)

When an intermediate BFR is processing a received MPLS packet, and one of the BFR's own BIER-MPLS labels rises to the top of the label stack, the BFR infers the sub-domain, SI, and BitStringLength from the label. The BFR then follows the forwarding procedures of [BIER_ARCH]. If it forwards a copy of the packet to a neighboring BFR, it first swaps the label at the top of the label stack with the BIER-MPLS label, advertised by that neighbor, that corresponds to the same SI, sub-domain, and BitStringLength. Note that when this swap operation is done, the TTL field of the BIER-MPLS label of the outgoing packet MUST be one less than the "incoming TTL" of the packet, as defined in Section 2.

Thus a BIER-encapsulated packet in an MPLS network consists of a packet that has:

- o An MPLS label stack with a BIER-MPLS label at the bottom of the stack.
- o A BIER header, as described in Section 3.
- o The payload.

The payload may be an IPv4 packet, an IPv6 packet, an ethernet frame, or an MPLS packet. If it is an MPLS packet, the BIER header is followed by a second MPLS label stack; this stack is separate from the stack that precedes the BIER header. For an example of an application where it is useful to carry an MPLS packet as the BIER payload, see [BIER_MVPN].

5. IANA Considerations

This document has no actions for IANA.

6. Security Considerations

As this document makes use of MPLS, it inherits any security considerations that apply to the use of the MPLS data plane.

As this document makes use of IGP extensions, it inherits any security considerations that apply to the IGP.

The security considerations of [BIER_ARCH] also apply.

7. Acknowledgements

The authors wish to thank Rajiv Asati, John Bettink, Nagendra Kumar, Christian Martin, Neale Ranns, Greg Shepherd, Ramji Vaithianathan, and Jeffrey Zhang for their ideas and contributions to this work.

8. Contributor Addresses

Below is a list of other contributing authors in alphabetical order:

Mach (Guoyi) Chen
Huawei

Email: mach.chen@huawei.com

Arkadiy Gulko
Thomson Reuters
195 Broadway
New York NY 10007
US

Email: arkadiy.gulko@thomsonreuters.com

Wim Henderickx
Alcatel-Lucent
Copernicuslaan 50
Antwerp 2018
BE

Email: wim.henderickx@alcatel-lucent.com

Martin Horneffer
Deutsche Telekom
Hammer Str. 216-226
Muenster 48153
DE

Email: Martin.Horneffer@telekom.de

Uwe Joorde
Deutsche Telekom
Hammer Str. 216-226
Muenster D-48153
DE

Email: Uwe.Joorde@telekom.de

Tony Przygienda
Ericsson
300 Holger Way
San Jose, CA 95134
US

Email: antoni.przygienda@ericsson.com

9. References

9.1. Normative References

- [BIER_ARCH] Wijnands, IJ., "Multicast using Bit Index Explicit Replication Architecture", internet-draft draft-wijnands-bier-architecture-02, December 2014.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.
- [RFC5331] Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream Label Assignment and Context-Specific Label Space", RFC 5331, August 2008.

9.2. Informative References

- [BIER_MVPN] Rosen, E., Ed., Sivakumar, M., Wijnands, IJ., Aldrin, S., Dolganow, A., and T. Przygienda, "Multicast VPN Using Bier", internet-draft draft-rosen-l3vpn-mvpn-bier-02, December 2014.
- [ISIS_BIER_EXTENSIONS] Przygienda, T., Ginsberg, L., Aldrin, S., and J. Zhang, "OSPF Extensions for Bit Index Explicit Replication", internet-draft draft-przygienda-bier-isis-ranges-01.txt, October 2014.
- [OSPF_BIER_EXTENSIONS] Psenak, P., Kumar, N., Wijnands, IJ., Dolganow, A., Przygienda, T., and J. Zhang, "OSPF Extensions for Bit Index Explicit Replication", internet-draft draft-psenak-ospf-bier-extensions-01.txt, October 2014.

Authors' Addresses

IJsbrand Wijnands (editor)
Cisco Systems, Inc.
De Kleetlaan 6a
Diegem 1831
BE

Email: ice@cisco.com

Eric C. Rosen (editor)
Juniper Networks, Inc.
10 Technology Park Drive
Westford, Massachusetts 01886
US

Email: erosen@juniper.net

Andrew Dolganow
Alcatel-Lucent
600 March Rd.
Ottawa, Ontario K2K 2E6
CA

Email: andrew.dolganow@alcatel-lucent.com

Jeff Tantsura
Ericsson
300 Holger Way
San Jose, California 95134
US

Email: jeff.tantsura@ericsson.com

Sam K Aldrin
Huawei Technologies
2330 Central Express Way
Santa Clara, California
US

Email: aldrin.ietf@gmail.com